

Reputation Formation and the Evolution of Cooperation in Anonymous Online Markets^{*}

Andreas Diekmann^a, Ben Jann^b, Wojtek Przepiorka^{cd}, Stefan Wehrli^a

^a ETH Zurich, Sociology, CLU D, Clausiusstrasse 50, CH-8092 Zurich, Switzerland

^b University of Bern, Institute of Sociology, Lerchenweg 36, CH-3000 Bern 9, Switzerland

^c Nuffield College, New Road, Oxford OX1 1NF, UK

^d University of Oxford, Department of Sociology, Manor Road, Oxford OX1 3UQ, UK

(forthcoming in *American Sociological Review*)

^{*} Version 24/07/2013. Word count excl. tables and figures: 11,153. We thank Stephanie Eckman, Debra Hevenstone, Claudia Jenny and six anonymous reviewers for their perceptive and helpful comments. This research was partly supported by the Swiss National Science Foundation (grant number 100017_124877). An online supplement to this article with additional information on previous research, data collection, and results is available from ASR's website. The data and analysis files used for this article can be downloaded at <http://www.socio.ethz.ch/research/datafiles>.

Author for correspondence: Wojtek Przepiorka (wojtek.przepiorka@sociology.ox.ac.uk)

Abstract: Theoretical propositions stressing the importance of trust, reciprocity and reputation for cooperation in social exchange relations are deeply rooted in classical sociological thought. Today's online markets provide a unique opportunity to test these theories using unobtrusive data. Our study investigates the mechanisms promoting cooperation in an online-auction market where most transactions can be conceived as one-time-only exchanges. We first give a systematic account of the theoretical arguments explaining the process of cooperative transactions. Then, using a large dataset comprising 14,627 mobile phone auctions and 339,517 DVD auctions, we test key hypotheses about the effects of traders' reputation on auction outcomes and traders' motives for leaving feedback. Our statistical analyses show that sellers with a better reputation have higher sales and obtain higher prices. Furthermore, we observe a high rate of participation in the feedback system, which is largely consistent with strong reciprocity – a predisposition to unconditionally reward (punish) one's interaction partner's cooperation (defection) – and altruism – a predisposition to increase one's own utility by elevating the interaction partner's utility. Our study demonstrates how strong reciprocity and altruism can mitigate the free-rider problem in the feedback system to create reputational incentives for mutually beneficial online trade.

Keywords: cooperation, trust, reciprocity, reputation, online market

1 Introduction

Social exchange has always been an important part of human sociality and arguably a driving force in the evolution of the human mind (Cosmides and Tooby 1992). In archaic societies, social exchange spanned a network of multiplex relations and was accompanied by elaborate ceremonies. A long tradition of research in anthropology and sociology on social exchange began with the work of Malinowski and Mauss. Ceremonial exchanges, such as the Kula (Malinowski 1922) or the potlatch (Mauss [1950] 1990), were formal rituals characterized by solemnity, decorum and disinterested generosity and occurred side by side with economic exchanges, which, in contrast, were guided by utilitarian motives (Heath 1976). Social exchange was governed by norms of reciprocity (Mauss [1950] 1990; Malinowski 1926; Gouldner 1960), and a failure to reciprocate was punished by the loss of reputation and status (Mauss [1950] 1990; Blau 1964).

With trade taking place over ever longer distances, markets transcended the borders of small communities, and there is historical evidence that social and economic institutions have evolved as a substitute for informal social control (Zucker 1986; Greif 1989). For example, trade in the early Middle Ages in Europe was characterized by geographical specialization, bookkeeping, and cashless payment. At that time, the Champagne fairs in France were a meeting point for traders from all over Europe. Milgrom, North, and Weingast (1990) discuss the emergence of a private adjudication system (the Law Merchant), which helped overcome trust problems among anonymous traders. Administered by private judges, it provided a platform for traders to settle disputes and document trading partners' dishonest behavior. Along with a system of notaries, such an institution could track and disseminate information about a trader's past behavior so that dishonest traders could be denied access to the market.

More recently, similar information-collecting institutions have emerged to overcome trust problems between money lenders and borrowers. The so-called credit bureaus function as information brokers that collect and collate information about borrowers' liabilities, credit histories (in particular arrears and defaults) and other characteristics (Japelli and Pagano 2002). Credit bureaus are run privately, often by a group of lenders, and the reporting of borrower data is voluntary. To overcome the free-rider problem in data reporting, the services provided by credit bureaus are based on reciprocity; only lenders who submit accurate information about their customers are granted access to a comprehensive customer database. This reciprocal information-sharing system creates financial incentives for lenders to contribute to the common good of customer information and for borrowers to maintain a good reputation by timely debt repayment. Japelli and Pagano (2002) show, moreover, that such information sharing is indeed associated with higher lending and lower default rates across 46 countries.

In a similar vein, cooperation in peer-to-peer online markets is sustained by an electronic reputation system that collects and disseminates information about traders' past behavior (Kollock 1999; Resnick et al. 2000). In online markets, tens of thousands of anonymous buyers and sellers trade with each other every day. These traders are not connected through a social network, in most cases interact only once (Resnick and Zeckhauser 2002) and – when cheating – are unlikely to be caught and fined by a central authority as many of these transactions transcend national borders (Przepiorka 2013). Crucial for the functioning of the reputation system is traders' willingness to submit feedback after transactions are finished. However, unlike in credit markets, information about traders' reputation is a collective good that is freely available to anyone (Bolton, Katok, and Ockenfels 2004). Since there is a cost for submitting feedback in terms of time and effort, traders have no real incentive to submit feedback and a second-order free-rider problem may exist (Heckathorn 1989). Assuming

strict rationality, actors would not rate their trading partners and, consequently, the whole market would degenerate or not evolve at all. How is it, though, that those online markets are so successful and enjoy increasing popularity?

Here we argue that other-regarding preferences such as altruism (Becker 1976) and strong reciprocity (Gintis 2000) play an important role in maintaining the provision of feedback in anonymous online markets. This feedback then generates the information necessary to create financial incentives for mutually beneficial trade. Based on a large sample of process-produced auction data obtained from eBay ($N \approx 350,000$) we first corroborate previous studies' findings that reputation has a market value, thus making it rational for traders to build and maintain a good reputation. We investigate the effect of reputation on the probability of sale and the selling price in a high- and a low-cost product market (mobile phones vs. DVDs respectively) and under high and low buyer uncertainty (used vs. new mobile phones respectively). Second, we investigate traders' motivation to leave feedback, without which reputational incentives would not accrue. Based on a large sample of timed rating events, we test partly competing hypotheses with regard to the possible motives behind traders' voluntary contributions to the feedback system. In our analyses, we take advantage of the repeated measures obtained for the same traders, to control for unobserved confounders (i.e., unobserved heterogeneity).

In what follows, we bring together different strands of research and describe theoretically as well as in a thorough empirical analysis the mechanisms that guide traders' cooperative behavior in an online market. In particular, we show how the institutional set-up of an online market engages traders' moral sentiments and material interests to create opportunities for mutually beneficial trade.

2 Theory and hypotheses

An interaction between a buyer and a seller in a typical online-auction market can be conceived as a process comprising four stages. First (the *initiation stage*), a seller decides to initiate an auction and sets some relevant parameters. Second (the *bidding stage*), potential buyers place bids and the highest bidder gets the offered item. Third (the *transaction stage*), the buyer pays the highest bid (i.e., selling price) and the seller ships the item. Finally (the *feedback stage*), the buyer can rate the seller and, depending on the implementation of the feedback mechanism, the seller can rate the buyer. Ratings can be positive or negative, are mostly accompanied by a short written comment and are immediately available to all market participants free of charge. In the following section, we outline our theoretical underpinning of and hypotheses about online traders' behavior at the transaction stage. Section 2.2 covers the feedback stage.

2.1 Transaction stage

The sequence of moves at the transaction stage is not predefined. However, since sellers cannot choose their buyers, most sellers demand advance payment and ship the item only after payment is received (Diekmann, Jann, and Wyder 2009). In this case, buyers cannot inspect the product before payment and a trust problem arises as sellers could deliver poor quality goods or keep the money without delivering at all (Akerlof 1970; Dasgupta 1988; Güth and Ockenfels 2003).¹ How do traders overcome this trust problem?

¹ Most economic and social exchange is sequential. An actor A transfers resources to an actor B and expects B to reciprocate at some point in the future so that both A and B earn the gains from trade. A trust problem arises when A cannot expect B to reciprocate with certainty (Coleman 1990: Ch. 5). We call A's expectation regarding B's propensity to reciprocate "trust". Thus, a trusting A is more likely to transfer resources to B. We call B's propensity to reciprocate "trustworthiness". B can be trustworthy because he or she is committed to act in A's interest out of self-interest (Hardin 2002), or because of other-regarding preferences and internalized norms of fairness and reciprocity (for other notions of trust see Cook, Hardin, and Levi 2007).

Evidence from laboratory experiments suggests that there may be a significant proportion of trustworthy sellers who care about the outcome of their trading partners because of other-regarding motives such as positive reciprocity, fairness considerations and altruism (Camerer 2003). However, these trustworthy sellers will be driven out of the market because they cannot keep up with the cheap offers of dishonest sellers. Bad experiences with dishonest sellers will decrease buyers' willingness to pay for a product and thereby reduce the trustworthy sellers' incentives to stay in the market (Akerlof 1970; Yamagishi et al. 2009). Alternatively, sellers may be trustworthy because they are embedded in dyadic relations or social networks (Granovetter 1992; Jones, Hesterly, and Borgatti 1997; Gulati and Gargiulo 1999; Buskens and Raub 2002), in which cooperation is enforced through direct and also through indirect reciprocity (Nowak and Sigmund 2005, see also Blau 1964) respectively. However, online consumer markets merely coordinate the supply and demand of end-product sellers and ultimate buyers; these markets are not characterized by customized and complex exchanges, and, as we will show below, repeated interactions between the same two traders are infrequent. Under these conditions, it is unlikely that network governance structures will emerge to solve potential exchange problems (Jones et al. 1997). Moreover, since online trade takes place across large geographic distances, it is also unlikely that offline commitment relations (Kollock 1994; Yamagishi, Cook, and Watabe 1998) or trust networks (DiMaggio and Louch 1998; Cook, Rice, and Gerbasi 2004) will form that would allow traders to obtain and disseminate information about their interaction partners. Finally, although online markets such as eBay try to establish norms of good conduct by providing discussion forums, chat rooms and clubs where traders have the opportunity to exchange information on selected topics, to our knowledge, there is no evidence that eBay traders have developed shared social norms of cooperation and, if they have, such norms will not prevent fraudulent traders from behaving opportunistically.²

² eBay Germany has 14.5 million active users and we counted 142 eBay clubs with a total of 2686 members

In anonymous online markets, the lack of social embeddedness is compensated for through an electronic rating system that disseminates information about traders' reputation in an efficient and systematic way (Resnick et al. 2000). From an organizational behavior perspective, an electronic reputation system is an exogenously established, fully connected network that embeds anonymous online traders in weak ties, through which information about their reputation is transmitted, without structurally limiting their choice of potential interaction partners and the pursuit of their goals (Uzzi 1997; Gulati and Gargiulo 1999). In other words, in an anonymous online market with a reputation system, traders can be conceived as structurally homogeneous and their interactions as isolated dyads (Granovetter 1992). The resulting social mechanism that resolves exchange problems is rather simple. With their reputation at stake, rational and self-regarding sellers who sufficiently care about their future business have a strong incentive to behave cooperatively. Even in an interaction with a buyer whom they are unlikely to meet again, these sellers will forgo the short-term temptation to abuse the buyer's trust in order to avoid a negative rating that would hamper future business. However, market entrants have no reputation and are therefore likely to be distrusted by buyers. Consequently, in order to enter the market, sellers with no feedback record have to allow prices to compensate buyers for the risk they take when buying from new sellers. If new sellers sufficiently care about future business, their good reputation will reimburse them for this initial investment in the future (Shapiro 1983; Resnick and Zeckhauser 2002; Przepiorka 2013).³ Thus, if buyers discriminate among sellers according to their reputation, selling prices

representing 0.019 percent of the active users (eBay.de, March 28, 2013). There are a somewhat higher number of participants in discussion forums but relative to millions of users and transactions, their proportion is negligible. Moreover, repeated transactions with the same trader are rare (see below). Thus, the density of the trader network will be close to zero, even for submarkets. Finally, online relations may be "too thin" (Hardin 2004) to establish norms of good conduct to the same extent as offline relations would (Granovetter 1992).

³ On eBay, traders can leave feedback only after a transaction, and for every transaction the market platform charges a small fee. Therefore, building a good reputation, even with sham transactions, requires time and money. This makes a good reputation costly to fake and the reputation system fairly robust against deliberate attempts to influence it. On platforms where the possibility to leave feedback is not restricted to the traders of a particular transaction, reputations can be faked more easily. Another aspect that stabilizes eBay's reputation

will be positively (negatively) correlated with a seller's number of positive (negative) ratings. Moreover, if the market does not clear because supply exceeds demand, an analogous hypothesis can be derived with respect to sellers' positive (negative) ratings and the probability of sale.

Previous studies have found a stronger absolute effect of negative ratings on outcomes than of positive ratings (Ba and Pavlou 2002). A possible explanation for this finding is that negative ratings carry more information because they express unmet expectations in relation to proper business practice while positive ratings merely confirm these expectations (Lucking-Reiley et al. 2007). Finally, with both new and used products a potential buyer may fear that the seller will not deliver after payment or that the quality of the item will be lower than advertised. However, in the case of used products, uncertainty is higher for the buyer because the description of the product condition is less standardized, leaving more room for dishonest behavior by sellers. Hence, a seller's reputation may be more important and therefore exhibit a larger effect than in the case of new products (Kollock 1994; Cook et al. 2004).

In summary, our hypotheses for the transaction stage are as follows:

- (H1) The number of positive (negative) ratings increases (decreases) the probability of a sale (H1a) and the selling price (H1b).
- (H2) The effects of negative ratings on probability of sale and selling price are stronger than the effects of positive ratings.
- (H3) The effects of positive and negative ratings are stronger for used products than for new products.

system is that the platform bans traders with too many negative evaluations and may prevent them from opening a new account under a different alias.

2.2 Feedback stage

We have argued in the previous section that the market value of a good reputation creates financial incentives for traders to behave cooperatively. However, rational and self-regarding traders will not provide feedback after a transaction is finished and thus reputational incentives will not accrue. Participation in the feedback system can be compared to voting, in that, just as with democratic elections, there is a collective good problem (Downs 1957). If citizens eschew the costs of voting, as is predicted by strict rationality principles, the democratic system collapses, as would the online market should no buyer provide feedback. Fortunately, however, a substantial proportion of citizens do vote and a large proportion of traders do leave feedback. Similar explanations can be given for both behaviors.

Consider a one-sided reputation system in which only buyers can rate sellers after a transaction. There are several possible explanations for buyers' participation in the feedback system. First, as suggested by signaling theory (e.g., Gambetta 2009), a buyer might leave feedback to gain a reputation for being an active rater. Such a reputation signals to the seller that the buyer will publicly comment on how a transaction plays out, which might enforce cooperative behavior on the side of the seller. Second, buyers might submit positive feedback because they are altruistic and the increase in the seller's utility elevates their own utility (Becker 1976).⁴ Since we assume a decreasing marginal utility of positive ratings, the likelihood of an altruistic buyer submitting feedback will decrease with an increasing seller score. Third, some buyers might leave feedback because they gain satisfaction from contributing to the public good provided by the feedback system. This idea corresponds to the notion of "warm glow" altruism (Andreoni 1990; see also Dellarocas, Fan, and Wood 2004). Finally, there is good reason to believe that some buyers are willing to punish defection or

⁴ Note that Becker's altruism model is agnostic regarding the emotional triggers of altruistic behavior. Regret and guilt, for instance, can be subsumed under the model.

reward cooperation without direct benefit. That is, giving feedback can be motivated by strong reciprocity (Gintis 2000; Fehr, Fischbacher, and Gächter 2002).⁵ Returning a favor with a favor and responding to misdeeds with sanctions is observed in experiments and seems to be rooted in human nature (de Quervain et al. 2004).

Until May 2007, eBay had a feedback system in which both transaction partners could leave feedback within 90 days. In such a two-sided rating system, in addition to the motives listed above, the feedback stage is open to strategic thinking. Although both buyers and sellers can be led by strategic (and non-strategic) motives, it is primarily sellers who have an interest in acting strategically at the feedback stage (see also Bolton, Greiner, and Ockenfels 2013). For example, the best way for a seller to establish a good reputation is to engage in good conduct, but there is always a small probability that a buyer is dissatisfied for reasons beyond the seller's control. To safeguard against a buyer's (possibly unjustified) negative response, a seller will prefer to hold up the threat of retaliation. Because a negative rating is more detrimental for sellers who have not yet established a good reputation than for sellers with many positive ratings, we expect sellers with fewer positive ratings to be more likely to strive for the second-mover position at the feedback stage. On the other hand, a seller could try to elicit positive feedback from a buyer by giving positive feedback first. Such a strategy makes sense if the seller believes that the buyer adheres to a norm of reciprocity (Gouldner 1960; Resnick and Zeckhauser 2002). If sellers make such an assertion, then the buyer's score will have a positive effect on the probability of the seller giving feedback first. A further variable likely to influence the seller's rating decision is whether the seller has already received a rating from the same buyer in a previous interaction. At the time of our data collection,

⁵ Fehr et al. (2002: 3) point out that the “essential feature of strong reciprocity is a willingness to sacrifice resources for rewarding fair and punishing unfair behavior *even if this is costly and provides neither present nor future material rewards for the reciprocator*” (original emphasis). Behavior as defined by strong reciprocity cannot therefore be explained by reciprocal altruism (Trivers 1971) or by reputation and indirect reciprocity (Nowak and Sigmund 2005).

repeated ratings from the same user did not affect the reputation score. Hence, a seller had no reason to elicit further positive feedback from a buyer by giving positive feedback first (see also Dellarocas et al. 2004).

For buyers, the non-strategic motives mentioned in the second paragraph of this section are still valid. In addition, if a buyer is motivated by strong reciprocity, positive feedback from the seller will further increase his or her likelihood of submitting a positive rating. In fact, it is hard to think of reasons why a self-regarding buyer would reciprocate a seller's positive feedback. Such behavior would be an indication of strong reciprocity at the feedback stage. Furthermore, indirect reciprocity (Nowak and Sigmund 2005), that is, rewarding sellers for their participation in the feedback system, might be another non-strategic motive in a two-sided system. According to indirect reciprocity, buyers will be more likely to give (positive) feedback to sellers with a good reputation.⁶ Finally, due to eBay's reputation system ignoring repeated ratings, buyers should be less likely to submit positive feedback if they provided a rating to the same seller in a previous interaction. If no such effect exists, altruistic motives that are based on the buyer's belief that the seller gains utility from receiving feedback (Becker 1976) can be ruled out. For our empirical analysis, we confine ourselves to the following hypotheses (see also Dellarocas et al. 2004):

- (H4) Reciprocity: If buyers and sellers are motivated by strong reciprocity at the feedback stage, their inclination to submit a positive (negative) rating increases after receiving positive (negative) feedback from the transaction partner.
- (H5) Effect of sellers' score on buyers' behavior⁷:

⁶ A high number of ratings can be interpreted as an indication of being an active rater due to the high correlation between incoming and outgoing ratings.

⁷ The two motives hypothesized under H5a and H5b are not mutually exclusive and may both be present simultaneously in a heterogeneous population. Empirical estimates, however, will provide information about which of the motives is predominant. Other motives such as the "warm glow" or signaling might also be present,

- (a) If buyers are motivated by Becker's altruism, the seller's score should have a negative effect on the probability of feedback submission.
 - (b) If buyers are motivated by indirect reciprocity, the seller's score should have a positive effect on their decision to provide feedback.
- (H6) Sellers' strategic behavior:
- (a) Sellers' inclination to make the first move at the feedback stage increases with their own score; and
 - (b) Sellers' inclination to make the first move at the feedback stage increases with the buyer's score.
- (H7) Previous interaction:
- (a) Sellers are less likely to submit positive feedback if they have already received a rating from the same buyer in a previous interaction.
 - (b) Buyers are less likely to submit positive feedback if they provided a rating to the same seller in a previous interaction.

3 Previous findings

Most studies investigating reputation effects in online markets analyze data collected from eBay, and most of these studies find evidence of a positive relation between sellers' reputation and prices or sales. Due to space limitations, we provide a review of these studies in the online supplement to this article (for earlier reviews, see Bajari and Hortacsu 2004; Resnick et al. 2006). Here, we instead focus on a number of important aspects in the ongoing discussion about reputation formation in anonymous online markets.

but we see no possibility to disentangle them with our data; these motives would be consistent with a null effect of sellers' score on buyers' behavior. Furthermore, effects as hypothesized in H5 could also be caused by fear of disagreeable e-mails by the seller if one does not leave feedback. However, we cannot offer any clear presumptions about how such fears would be related to sellers' score.

In their review of 15 empirical eBay studies, Resnick et al. (2006) point out that the quasi-experimental designs of these studies cannot account for unobserved heterogeneity in a satisfactory way. Thus, if unobserved factors such as the pre-purchase communication between buyers and sellers or the quality of the product description are correlated with sellers' reputation, then the estimates of the regression coefficients can be biased. In our data set, we have repeated measures for a considerable proportion of both buyers and sellers. This allows us to cope with problems of unobserved heterogeneity by using fixed-effects regressions (Allison 2009).

It has been argued that most eBay studies claim to estimate buyers' willingness to pay for sellers' reputation but, in fact, only estimate what buyers do pay, which is a poor estimate of their willingness to pay (Snijders and Weesie 2009). The argument is that the highest bids may depend on many other factors not observable in the data obtained from eBay. For instance, bidding is a dynamic process and potential buyers' bidding decisions also depend on their beliefs about other buyers' reservation prices. Moreover, potential buyers could consider several offers simultaneously, but their initial choice sets are unobserved or difficult to define (although see Livingston, 2003). In our study, we do not claim to be estimating buyers' willingness to pay for sellers' reputation. For the sake of our argument, it is enough to show that buyers, on average, pay more for sellers with a higher reputation and thus create financial incentives for sellers to behave cooperatively in the market.

The Achilles' heel of a reputation system remains market participants' willingness to rate each other. If participants do not provide truthful ratings, the market loses its capacity to identify fraudulent traders and reputation loses its value. For buyers and sellers on eBay, Resnick and Zeckhauser (2002) report a rating frequency of 52% and 61%, respectively; Dellarocas et al. (2004) find 68% and 78%; and Jian, MacKie-Mason, and Resnick (2010)

report rates of 33% and 55%. Moreover, averaging across seven different countries' eBay sites, Bolton et al. (2013) report a rating frequency of 71%. They attribute the high rating frequency to the two-sided rating system implemented on eBay. On the one hand, a two-sided system provides incentives to leave feedback because of positive reciprocity expectations. On the other hand, the threat of retaliation against a negative rating leads to a positive evaluation bias. Based on a data set of ratings obtained from eBay, Dellarocas and Wood (2008) also identify reciprocity as an important factor driving behavior at the feedback stage. Moreover, they suggest that not leaving feedback indicates dissatisfaction and, based on the information obtained from the observed rating patterns, they estimate the proportion of good, bad, and mediocre transactions. They show that the 99% positive rating usually observed on eBay is a biased estimate of the proportion of good transactions; accounting for the "sound of silence" (i.e., transactions where one or both traders did not leave feedback) gives an estimate of 79% and 86% for satisfied buyers and sellers respectively. In a similar vein, Jian et al. (2010) develop a model to estimate the proportion of different rating strategies from eBay feedback data. The authors assume three rating strategies (*do not give feedback*, *give feedback unconditionally*, *reciprocate feedback*) and estimate both buyers' and sellers' propensity to choose one of the strategies. They find that the reciprocal feedback strategy is chosen in 23% or 20% of the cases, whereas the unconditional feedback strategy is chosen in 38% or 47% of the cases by a median buyer or seller respectively.

In our study, we complement these results by analyzing time-specific feedback decisions. Whether or not the trading partner has provided feedback first is an explanatory variable in our analysis. We estimate discrete-time event-history models, an approach which does not rely on restrictive assumptions about the timing distribution of traders' feedback. We also account for unobserved heterogeneity by including buyer and seller fixed effects in our models. Note, however, that what we observe in our data is behavior and, therefore, any

inference we make about traders' underlying motives is based on the consistency of the behavioral data with our theoretical propositions.

4 Data

Our data, collected from the German eBay platform, contains information on auctions of mobile phones and DVDs that ended between December 1, 2004 and January 7, 2005 (for details on data collection, see the online supplement). We restrict our analysis to offers from sellers in Germany and exclude private auctions, where the buyer is unknown, as well as auctions in which multiple pieces are offered that can be purchased by multiple buyers. Some auctions ended early because the seller withdrew the auction before the first bid, the seller stopped the auction with the current bid as the winning bid or the auction ended with a "Buy It Now" (i.e., fixed price) purchase. We exclude these cases from our analysis. Finally, minor inconsistencies in the regular expressions caused missing or ambiguous data on key variables for a small fraction of auctions, which we also exclude (in total, 14.5% of mobile phones and 10.4% of DVDs were excluded for the above reasons).

To estimate reputation effects on the highest bids in auctions, it is sensible to select a homogeneous good to avoid bias due to unobserved heterogeneity (Diekmann et al. 2009). To ensure homogeneity in our mobile phone data, we manually sorted out offers that contained complementary goods or multiple items (23.1%) and divided the sample into new (unused and in original packaging) and used products. The total sample therefore includes 14,627 mobile phones of seven different models, 37.6% of which are new and 95.7% of which received at least one bid (i.e., their auctions were successful).⁸ In our data, the mean selling price was 217 Euros for new mobile phones and 150 Euros for used mobile phones (see Table 1 containing

⁸ We collected data for five more models, but we exclude these cases from our analysis because the models were released prior to 2004. The excluded cases are comprised almost entirely of used products and thus are more heterogeneous than the offers of newer models. Including these older models in our analysis does not change our main findings.

descriptive statistics of our key variables; additional variables are documented in Table A1 in the online supplement). To evaluate whether reputation effects can also be observed for low-cost products, we analyze a sample of 339,517 DVD auctions, 53.3% of which were successful. We collected data on all DVD auctions in the mentioned time span, covering hundreds of different movies. Thus, product homogeneity cannot be assumed. Also, due to the lack of a reliable automatic categorization method, we do not distinguish between new and used DVDs. Moreover, unlike the mobile phone market, with 1.25 auctions per seller on average, the DVD market, with an average of 10.24 auctions per seller, is dominated by a relatively small number of professional sellers.

We operationalize buyers' and sellers' reputation with their number of positive and negative ratings. Since the distributions of these variables are highly right-skewed, we take the logarithm of their values (after adding 1). Note that in terms of ratings too, sellers in the DVD market are considerably larger than sellers in the mobile phone market.

[Table 1 about here.]

At the feedback stage, we have the time (in days) until feedback for both the seller and the buyer of a successful auction. Recall that these variables are censored at 90 days. The mean duration seems to be lower in the DVD market, and positive feedback is very common, whereas neutral and negative ratings are rare. Negative ratings are more frequent in auctions with mobile phones than DVDs, indicating that problematic transactions are more likely when they involve more complex and more expensive products (although this could also be due to the lower degree of professionalism in the mobile phone market). Another important variable at the feedback stage is whether the seller or buyer has already received a rating from the

same person in a previous interaction. Such repeated encounters are more likely in the DVD market (5%) than in the mobile phone market (3%).⁹

5 Results

5.1 The effect of reputation on auction success

In this section, we report and discuss reputation effects on sales and selling price. Table 2 gives an overview of the results for our key variables. Although we control for many additional variables in our model estimations, due to space limitations, we report and discuss their effects in the online supplement only.

Models M1 through M3 in Table 2 show the results from logistic regressions of auction success on sellers' positive and negative ratings. In line with our hypothesis (H1), we find a significantly positive effect of positive ratings and a negative effect of negative ratings on the probability of sale for new mobile phones (model M1) and for DVDs (M3). Contrary to our expectation, we find no significant effects for used mobile phones (M2). Recall, however, that more than 95% of all mobile phones were sold, which leaves relatively little information for the estimation of the reputation effects on the probability of sale.

[Table 2 about here.]

Since we are dealing with auctions and not with fixed price offers, a large part of the reputation effect will be exerted on price. Table 2 also reports results from linear regressions of selling price on sellers' positive and negative ratings (models M4 through M6). As expected, we observe a strong relation between a seller's reputation and the highest bid in our

⁹ These numbers are calculated based on traders' entire rating histories. They are lower bounds for the actual proportions of repeated encounters because transactions appear in the feedback history only if a rating occurred.

data. For both mobile phones and DVDs, the number of positive ratings has a significantly positive effect on price, and negative ratings reduce it.¹⁰ Since both the number of ratings and the selling price are log-transformed, the estimated coefficients can be interpreted as a percent increase in the selling price due to a one percent increase in the number of ratings (the elasticity). Thus, doubling the number of positive ratings increases the selling price of new mobile phones (M4) by about $0.005 \times \ln(2) \times 100 = 0.35\%$. For used mobile phones (M5) and DVDs (M6), the effect of a 100% increase in positive ratings is about 0.55% and 3.7%, respectively. Evaluating these relative effects at the mean prices for new and used mobile phones and DVDs (see Table 1) yields absolute effects of 61, 97, and 27 cents, respectively. While these amounts may appear small at first sight, they are not, in fact, if one takes into account the range of sellers' reputations in our sample, where differences of 1000% (e.g., 40 vs. 440 positive ratings) and more are the rule rather than an exception. As predicted in H2, the absolute effect of negative ratings on price and sales is larger than the effect of positive ratings in all models (M1 through M6), although the difference is statistically significant at the 5% level only in model M4 ($p = 0.003$ for a two-sided test). In terms of absolute effects at mean prices, doubling the number of negative ratings decreases the price by 1.3 Euros, 2.2 Euros, and 52 cents in models M4, M5 and M6, respectively.

In line with our expectation that used products exhibit more uncertainty than new products and that therefore buyers will pay a higher reputation premium to sellers of used products (H3), the effects of positive and negative ratings on the selling price are stronger for used mobile phones (M5) than for new mobile phones (M6).¹¹ However, recall that only for new

¹⁰ Censoring at the starting price might bias the effects of positive and negative ratings on price because our estimation is based on successful auctions only. To evaluate the robustness of our results, we applied various models that take censoring into account (such as censored normal regression and Heckman's sample selection model). Results are very similar in all models (not shown).

¹¹ A bootstrap test of the joint null-hypothesis that the absolute coefficients of positive and negative ratings are not larger in the model for used mobile phones than in the model for new mobile phones yields a p -value of 0.04. This p -value is an *empirical strength probability* computed as the fraction of bootstrap samples in which the

mobile phones does reputation also have a significant effect on sales. If we take into account both outcomes (the selling price and the probability of sale), the total effect of reputation is even larger for new than for used mobile phones, although the difference is statistically insignificant (not shown).¹²

As mentioned above, we cannot assume product homogeneity for our DVD dataset. We therefore replicate our analysis using a model with fixed effects for movie titles (M9).¹³ Although smaller in absolute size than in the ordinary DVD model, the effects of positive and negative ratings are still highly significant.

One could object that our effects are spurious because of unobserved seller heterogeneity. For example, some sellers might be more successful for reasons unrelated to reputation and, therefore, more likely to have a good reputation, resulting in a spurious correlation between reputation and selling prices. We can check this argument by analyzing the effects of changes in reputation for sellers who appear repeatedly in our dataset. We estimate models with seller fixed effects for new mobile phones and used mobile phones (M7 and M8) and a model with crossed fixed effects for sellers and titles for DVDs (M10).¹⁴ In the model for new mobile phones, the effects of positive and negative ratings persist, as does the effect of negative ratings in the model for DVDs, but for used mobile phones, the effects disappear. However, the fact that we still find reputation effects in most of our models provides strong support for the reputation hypothesis (H1). Note that accounting for seller fixed effects results in fairly

point estimates for the coefficients are in accordance with the null-hypothesis (see Davison, Hinkley, and Young 2003: 151–152). The p -values for the separate hypotheses are 0.13 for positive and 0.10 for negative ratings.

¹² We determine the total effect by computing average marginal effects on the logarithm of the expected price, where the expected price is computed as the predicted selling price weighted by the estimated probability of sale.

¹³ Title matches could not be found for all auctions, even after removing special characters. Therefore, the sample size in this analysis is reduced to 113,276 sold items. These items were sold by 15,545 different sellers and comprised 18,054 different titles with 6.3 auctions per title on average (the largest group being *Harry Potter and the Prisoner of Azkaban*, with 383 auctions).

¹⁴ For the two-way fixed effects model, we use Schmiedler's (2009) implementation of an algorithm proposed by Guimarães and Portugal (2010).

conservative estimates of the reputation effects. The estimates are only based on changes in sellers' reputations during the observation window of about one month, while all between-seller variability is discarded. Consequently, the model estimations will be based on sellers with high trading frequency, which are more likely to have a good reputation, while newcomers, for whom an increase in reputation might be most valuable, tend to be excluded.

In sum, our analyses replicate previous findings of a positive (negative) effect of positive (negative) ratings on sales and prices (H1) and we also find a larger absolute effect of negative ratings as compared to positive ratings (H2). Our additional model estimations also make it rather unlikely that the effects we find are spurious. However, the evidence for the uncertainty hypothesis (H3) is mixed; the reputation effect does not seem to be larger for used mobiles than for new ones.

5.2 Participation in the feedback system

As argued above, if too few traders participate in the feedback forum, the reputation system will be ineffective in detecting cheaters, and reputational incentives for (first-order) cooperation at the transaction stage will not accrue. However, we do find high levels of (second-order) cooperation at the feedback stage. Buyers submitted positive feedback in over 65% of all transactions in our mobile phone sample, whereas only 3% of the transactions were rated neutrally or negatively. In the DVD sample, the difference is even larger. Buyers rated 83% of the transactions positively; neutral or negative feedback was submitted in less than 1% (see Table 1). Furthermore, Table 3 shows that more than 80% of the transactions in the mobile phone market and 90% of the DVD transactions were rated by at least one trader.

[Table 3 about here.]

After an auction has ended, traders have 90 days to submit their feedback, and as traders can observe their partners giving them feedback, their feedback decisions are not independent from each other. Therefore, it is important to take the exact timing of traders' feedback decisions into account when analyzing their rating behavior. Figure 1 gives an overview of the distribution of feedback decisions over the 90 days. The survival functions (upper panel) show how the proportions of buyers and sellers who have not yet given feedback decrease over time; the hazard rates for positive and negative ratings (middle and lower panel respectively) show, roughly speaking, the feedback probability on a given day conditional on no feedback had been given before; dotted lines indicate 95% confidence intervals (see the online supplement for computational details).

While survival curves for sellers and buyers are synchronized in both markets, ratings are submitted faster in the DVD market. After 10 days, almost 50% of sellers and buyers had submitted a rating in the DVD market, but only about 40% had done so in the mobile phone market. Most notably, sellers are somewhat more likely to quickly submit a positive rating. That is, the hazard rate of a positive rating is higher for sellers than for buyers over the first couple of days in both markets. This indicates that some sellers might try to elicit a positive rating from the buyer by taking the first step. For negative ratings, the picture is different (note that the scale of the hazard rate is magnitudes smaller than for positive ratings). The curves are very similar for sellers and buyers, although in the mobile phone market the hazard rate for buyers is higher than for sellers. Furthermore, negative ratings are submitted more slowly than positive ratings. The hazard rates for positive ratings peak between days five and 10, while they are highest between days 15 and 35 for negative ratings. Interestingly, unlike for positive ratings, we observe an increase in the hazard rates of negative ratings between days 80 and 90 (although confidence intervals tend to be large as data mass gets small in this region). This indicates that some actors may delay their negative rating to escape retaliation.

[Figure 1 about here.]

The survival curves and hazard rates in Figure 1 only provide descriptive information on the distribution of ratings over time. To analyze the effects of covariates we divide the process time into days and approximate the hazard rates using discrete-time linear probability models containing indicator variables for different segments of the process time (i.e., we model the baseline hazard rates shown in Figure 1 as step functions).¹⁵ We also include two-way buyer and seller fixed effects in the models to control for unobserved buyer and seller characteristics that might otherwise bias the results. For instance, a positive relation between traders' reputation and their rating probability could be due to the fact that frequent raters accumulate more ratings because their ratings are reciprocated. Likewise, transaction quality or speed might be higher among traders with better reputations, thus increasing the likelihood of receiving positive feedback or decreasing the time-to-feedback reaction. Such potential biases are kept at a minimum by including seller and buyer fixed effects in the models.

Table 4 provides the estimates from the discrete-time linear probability models for positive and negative feedback from buyers and sellers in the DVD market (we do not present results for mobile phones here because the sample is too small for the two-way fixed effects estimation). The displayed coefficients have been scaled by a factor of 100; they can be interpreted as percentage-point effects on the conditional probability of submitting a rating on a specific day given that no rating had yet been submitted.

[Table 4 about here.]

¹⁵ The segments are chosen such that in each interval a sufficient number of rating events is available. For positive feedback we use one-day intervals up to day 30, followed by two-day intervals up to day 50, followed by five-day intervals. For negative feedback the intervals in days are: 1–10, 11–15, 16–20, 21–25, 26–30, 31–40, 41–50, 51–70 and 71–90. The exact choice of the intervals does not seem to matter much for our findings.

Our results are consistent with the hypothesis that ratings are driven by strong reciprocity (H4). All models in Table 4 include time-dependent variables indicating whether the trading partner submitted a positive, neutral, or negative rating first. Once a seller submits a positive rating, the buyer's conditional probability of submitting a positive rating sharply increases by about five percentage points (M11); for sellers, this effect is even stronger (M12). Likewise, the probability that a negative rating will be reciprocated with a negative rating is very high (M13 and M14). Note that these effects are much smaller because of the low baseline probability.¹⁶ A somewhat inconclusive finding is the positive effect of a buyer's negative rating on the seller's probability of giving a positive rating.

Moreover, our results support hypothesis H5a as opposed to H5b. The more positive ratings a seller has, the less likely the buyer is to submit feedback. That is, in addition to strong reciprocity, buyers' rating behavior is consistent with Becker's altruism rather than with warm glow altruism, signaling, or indirect reciprocity. Also, the remaining evidence in Table 4 is consistent with the view that buyers, in particular, take their interaction partners' utility into account when leaving feedback. First, in line with our hypothesis H7b, buyers (and sellers) seem to acknowledge that traders to whom they provided a rating before do not benefit from another rating. Second, buyers are more likely to leave a positive rating the more negative ratings the seller has. This might reflect buyers' willingness to support sellers who want to redeem their good reputation. Finally, we find that buyers (and sellers) are reluctant to give a negative rating if a trading partner already has many negative ratings, but they are more likely to give a negative rating to a trading partner with many positive ratings. This might reflect the

¹⁶ In the subsample used for the two-way fixed effects estimation, the average conditional probability of feedback being given on a specific day is about 3% for positive feedback and 0.01% for negative feedback.

fact that, after a failed transaction, traders' disappointment with a trading partner with a good reputation is higher than with a partner with a bad reputation.

We do not find support for our hypotheses regarding sellers' strategic behavior at the feedback stage. Neither the sellers' own number of positive ratings (H6a) nor the buyers' number of positive ratings (H6b) exhibits a positive effect on the sellers' probability of leaving positive feedback. Also, sellers are not less likely to give feedback to a buyer from whom they have received a rating in a previous interaction, which contradicts our hypothesis H7a.

In sum, we observe high levels of participation in the feedback system, which seem to be mainly driven by strong reciprocity and altruism. Both sellers and buyers reciprocate received ratings from their transaction partners (H4) and appear to take their interaction partners' utility into account in their feedback decisions (H5a, H7b). We do not find evidence for sellers' strategic considerations, as our corresponding hypotheses cannot be confirmed (H6a, H6b, H7a). However, this does not mean that strategic motives are unimportant or do not exist; according to our descriptive results, sellers are more likely to submit a fast positive feedback, which can be seen as a strategic attempt to elicit reciprocation.

6 Discussion and conclusions

In commerce, trust problems have been ubiquitous ever since humans started trading commodities, and cultural evolution together with the creativity of "social engineers" led to a great variety of institutions that helped to mitigate these trust problems. The reputation system of the Maghribi traders in the Middle Ages described by Greif (1989), the "law merchant" system analyzed by Milgrom et al. (1990), the formation of formal social structures in nineteenth-century North America explored by Zucker (1986), and informal network

governance structures as described by Jones et al. (1997) are examples of institutions that evolved to overcome the uncertainty inherent in economic transactions. Factors such as the degree of social embeddedness of the traders (Granovetter 1992; Di Maggio and Louch 1998; Buskens and Raub 2002), the type of commodity (Kollock 1994), the complexity of the interaction (Jones et al. 1997) and, most importantly, the absence of timely and inexpensive provision of information about one's trading partners' reputation (Milgrom et al. 1990) have been identified as important determinants of their evolution.

Recent technological progress has drastically diminished the cost of fast information transmission, and auction markets were among the first to establish online reputation systems for economic transactions. Based on a seller's reputation, buyers can trade off prices against the possibility of being cheated, and the price premium for a good reputation provides a financial incentive for sellers' cooperative behavior. What is more, traders do not have to be embedded in offline social networks to access information about other traders as long as they have access to the market platform online. In the last decade, online reputation systems have spread all over the web, and the technical possibilities of the Internet have brought their functioning close to perfection. However, the existence and stability of online reputation systems cannot be taken for granted. A reputation system's effectiveness in promoting (first-order) cooperation crucially depends on traders' voluntary provision of feedback. As a trader gains no direct benefits from leaving feedback, participation in the feedback system is subject to a (second-order) free-rider problem.

Our study investigates the role of reputation in promoting cooperation and the process of reputation formation in the field. That is, we explore the mechanisms underlying cooperation by observing transactions and ratings among traders in anonymous online markets in an unobtrusive way. We consider online markets as integrated systems of institutional rules that

govern traders' decisions, and our analysis covers the entire chain of interactions between a buyer and a seller, starting with the seller's offer and ending with either or both leaving feedback. Our theoretical considerations are rooted in classical sociology and anthropology, with their insights into the roles of trust, reciprocity and reputation in social exchange, but we also draw on recent findings from experimental social sciences. Based on a large sample of data from the mobile phone and DVD markets, we test various and partly competing hypotheses derived from our theoretical argument.

In both markets, we find that sellers' number of positive ratings has a positive effect on sales and selling prices, whereas sellers' negative ratings have the opposite effect (H1). Moreover, in line with our hypothesis (H2), negative ratings affect sales and prices more than positive ratings. This corroborates previous findings that reputation has a market value and that buyers take into account information about sellers' reputation when they place their bids (see the online supplement for a review of previous findings). However, the empirical evidence for the uncertainty hypothesis (H3) is mixed. As hypothesized, reputation has a larger effect on price for used mobile phones than for new ones, but this is not the case for the probability of sale; the gross effect of reputation, taking into account both the price and the probability of sale, seems even larger for new mobile phones than for used ones. A possible explanation for this finding is that used products attract buyers who are willing to take a risk and therefore do not take information about sellers' reputation into account as carefully as buyers of new products do.

At the feedback stage, cooperation is very high. Sixty percent of transactions in the mobile phone market and 80% in the DVD market are rated by both sellers and buyers, and only for a small fraction of transactions is there no rating at all. This is comparable to what Dellarocas et al. (2004) and Bolton et al. (2013) find in their eBay data. Our study corroborates, moreover,

that buyers' and sellers' behavior at the feedback stage is largely consistent with strong reciprocity (H4). Although there are no apparent benefits from reciprocating a rating received from one's trading partner, the likelihood of providing similar feedback sharply increases for both buyers and sellers after receiving a rating. The effects are strong and highly significant for both positive and negative ratings. This is in line with the statistics reported in Bolton et al. (2013) and with what Dellarocas and Wood (2008) and Jian et al. (2010) find based on their model estimations.

Besides reciprocity, altruistic preferences seem to motivate traders' (in particular, buyers') decisions. Buyers are more likely to give a positive rating to traders with few positive or many negative ratings (H5a). This contradicts Jian et al. (2010), who find that both buyers and sellers are more likely to give feedback to experienced traders than to inexperienced traders. We argue that this could be an artifact resulting from Jian et al. (2010) disregarding transaction quality in their model estimations. Transaction quality is likely to be higher for experienced traders and could thus lead to a positive correlation between traders' experience and the probability of giving feedback. We circumvent this problem by including partner fixed effects in our model estimations; if we do not include them, our results are similar to those of Jian et al. (2010). Moreover, we find that buyers (and sellers) are less likely to give positive feedback to the same trading partner repeatedly (H7b). Both H5a and H7b are in line with Becker's (1976) notion of altruism; our results are consistent with the idea that traders take their interaction partners' utility into account when leaving feedback and are more likely to give feedback the more beneficial it is for their partner.

Contrary to our expectations and unlike the results by Dellarocas et al. (2004), we find no evidence for sellers' strategic considerations at the feedback stage. In particular, we find no support for the hypotheses that sellers strategically take into account their own or the buyers'

score at the feedback stage (H6a and H6b). Dellarocas et al.'s (2004) conclusion was based on the finding that a buyer's first move at the feedback stage had a negative effect on the seller's overall probability of leaving feedback. We argue that this result could be an artifact resulting from their using a probit model; if sellers have a lower *a priori* probability of providing a rating, buyers will more likely be the first movers at the feedback stage. The event-history models used in our analysis do not suffer from this problem because the buyer's first move is treated as a time-varying variable. Yet, traders seem to anticipate their trading partners' reciprocal behavior, especially in the case of negative feedback, as some try to avoid retaliation by submitting their negative rating close to the end of the rating period. Thus, we cannot conclude that strategic motives are unimportant in traders' decisions to leave feedback; they are difficult to identify in our data.

In Gouldner's (1960: 176) words, the norm of reciprocity is a “‘starting mechanism’ [that] helps to initiate social interaction”, and it is also a cornerstone in Blau's (1964) social exchange theory. Both authors elaborate on the principle of reciprocity and outline its implications for the emergence and stability of exchange relations. Their notion of reciprocity and the reciprocity we observe in our data is similar to the notion of “strong reciprocity” pioneered by Gintis (2000). According to strong reciprocity, actors will return a favor with a favor and retaliate against an unfriendly act without expecting to be compensated for the costs they incur in doing so. In laboratory experiments, Fehr et al. (2002) were able to demonstrate that strong reciprocity is a key mechanism in promoting cooperation in voluntary contribution games and sparked a cross-disciplinary debate concerning the determinants of human social cooperation (Guala 2012). In our study, we show how the interplay of simple institutional rules and human motivation establishes a high level of cooperation among anonymous traders in online markets. Based on our findings we argue that, in these markets, altruism and strong reciprocity are likely to play an important role in maintaining (second-order) cooperation at

the feedback stage, which thus generates the information necessary to create financial incentives for (first-order) cooperation at the transaction stage. In other words, giving positive or negative feedback – a second-order issue – can be conceived as rewarding or punishing a trading partner’s first-order cooperation or defection respectively (Resnick and Zeckhauser 2002). Given that the constrained set-up of online markets mirrors laboratory conditions relatively well, our findings bolster the external validity of experimental evidence of social cooperation and peer punishment. At the same time, our study makes apparent that unobtrusive behavioral data from online platforms need to be complemented with laboratory and field experiments, which are better suited to disentangling the various motives behind human behavior. Although strong reciprocity and altruism appear as the most likely motives behind the behavioral patterns we observe at the feedback stage, carefully designed experiments should be used to verify our results and rule out alternative explanations.

Electronic reputation systems have been constantly refined to better meet the requirements of a steadily growing online community. In fact, the change in eBay’s reputation system in spring 2007 was guided by a thorough theoretical and empirical analysis which combined field data from online markets with evidence from laboratory experiments (Bolton et al. 2013). The main concern with the old two-sided feedback system was that it might foster a positive evaluation bias. On the one hand, it encourages positive evaluations due to positive reciprocity expectations and, on the other hand, it deters truthful negative evaluations due to negative reciprocity expectations. Without abandoning the old system, buyers were newly given the opportunity to leave a more detailed and anonymous seller rating. That is, after a transaction, buyers could rate sellers on several attributes such as item description, communication, transaction speed and costs, etc., and sellers were not able to identify the buyer who rated them as they were only shown average feedback scores on each attribute. Surprisingly, although the costs for leaving feedback increased for buyers, both buyers’ and

sellers' conventional rating behavior hardly changed ten weeks into the system change and 70% of buyers who gave conventional feedback also gave detailed seller feedback (Bolton et al. 2013). This provides additional support for our assertion that cooperation at the feedback stage is motivated by strong reciprocity and altruism, rather than strategic considerations encouraged by the two-sided feedback system.

References

- Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics* 84:488–500.
- Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.
- Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal* 100:464–477.
- Ba, Sulin and Paul A. Pavlou. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior." *MIS Quarterly* 26:243–268.
- Bajari, Patrick and Ali Hortacsu. 2004. "Economic Insights from Internet Auctions." *Journal of Economic Literature* 42:457–486.
- Becker, Gary S. 1976. "Altruism, Egoism and Genetic Fitness: Economics and Sociobiology." *Journal of Economic Literature* 14:817–826.
- Blau, Peter M. 1964. *Exchange and Power in Social Life*. New York: Wiley.
- Bolton, Gary E., Ben Greiner, and Axel Ockenfels. 2013. "Engineering Trust: Reciprocity in the Production of Reputation Information." *Management Science* 59:265–285.
- Bolton, Gary E., Elena Katok, and Axel Ockenfels. 2004. "How Effective are Electronic Reputation Mechanisms? An Experimental Investigation." *Management Science* 50:1587–1602.
- Buskens, Vincent and Werner Raub. 2002. "Embedded Trust: Control and Learning." Pp. 167–202 in *Group Cohesion, Trust and Solidarity*. Vol. 19, *Advances in Group Processes*, edited by E. J. Lawler and S. R. Thye. Amsterdam: JAI.
- Camerer, Colin F. 2003. *Behavioral Game Theory*. Princeton (NJ): Princeton University Press.
- Coleman, James S. 1990. *Foundations of Social Theory*. Cambridge (MA): The Belknap Press of Harvard University Press.
- Cook, Karen S., Russell Hardin, and Margaret Levi. 2007. *Cooperation Without Trust?* New York: Russell Sage Foundation.
- Cook, Karen S., Eric R. W. Rice, and Alexandra Gerbasi. 2004. "The Emergence of Trust Networks under Uncertainty: The Case of Transitional Economies—Insights from Social Psychological Research." Pp. 193–212 in *Creating Social Trust in Post-Socialist Transition*, edited by J. Kornai, B. Rothstein, and S. Rose-Ackerman. New York: Palgrave MacMillan.
- Cosmides, Leda and John Tooby. 1992. "Cognitive Adaptations for Social Exchange." Pp. 163–228 in *The Adapted Mind*, edited by J. H. Barkow, L. Cosmides, and J. Tooby. New York: Oxford University Press.
- Dasgupta, Partha. 1988. "Trust as a Commodity." Pp. 49–72 in *Trust: Making and Breaking Cooperative Relations*, edited by D. Gambetta. Oxford: Basil Blackwell.
- Davison, A. C., D. V. Hinkley, and G. A. Young. 2003. "Recent Developments in Bootstrap Methodology." *Statistical Science* 18:141–157.
- Dellarocas, Chrysanthos, Ming Fan, and Charles A. Wood. 2004. "Self-Interest, Reciprocity, and Participation in Online Reputation Systems." Working Paper 205, MIT Center for Digital Business.
- Dellarocas, Chrysanthos and Charles A. Wood. 2008. "The Sound of Silence in Online Feedback: Estimating Trading Risks in the Presence of Reporting Bias." *Management Science* 54:460–476.
- Diekmann, Andreas, Ben Jann, and David Wyder. 2009. "Trust and Reputation in Internet Auctions." Pp. 139–165 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage.

- DiMaggio, Paul and Hugh Louch. 1998. "Socially Embedded Consumer Transactions: For What Kinds of Purchases Do People Most Often use Networks?" *American Sociological Review* 63:619–637.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Row.
- Fehr, Ernst, Urs Fischbacher, and Simon Gächter. 2002. "Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms." *Human Nature* 13:1–25.
- Gambetta, Diego. 2009. "Signaling." Pp. 168–194 in *The Oxford Handbook of Analytical Sociology*, edited by P. Hedström and P. Bearman. Oxford: Oxford University Press.
- Gintis, Herbert. 2000. "Strong Reciprocity and Human Sociality." *Journal of Theoretical Biology* 206:169–179.
- Gouldner, Alvin W. 1960. "The Norm of Reciprocity: A Preliminary Statement." *American Journal of Sociology* 25:161–178.
- Granovetter, Mark. 1992. "Problems of Explanation in Economic Sociology." Pp. 25–56 in *Networks and Organizations: Structure, Form, and Action*, edited by N. Nohira and R. G. Eccles. Boston: Harvard Business School Press.
- Greif, Avner. 1989. "Reputation and Coalitions in Medieval Trade: Evidence on the Maghribi Traders." *Journal of Economic History* 49:857–882.
- Guala, Francesco. 2012. "Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate." *Behavioral and brain sciences* 35:1–15.
- Guimarães, Paulo and Pedro Portugal. 2009. "A Simple Feasible Alternative Procedure to Estimate Models with High-Dimensional Fixed Effects." *Stata Journal* 10:628–649.
- Gulati, Ranjay and Martin Gargiulo. 1999. "Where Do Interorganizational Networks Come From?" *American Journal of Sociology* 104:1439–1493.
- Güth, Werner and Axel Ockenfels. 2003. "The Coevolution of Trust and Institutions in Anonymous and Non-anonymous Communities." Pp. 157–174 in *Jahrbuch für Neue Politische Ökonomie*, edited by M.J. Holler, H. Kliemt, D. Schmidchen, and M. Streit. Tübingen: Mohr Siebeck.
- Hardin, Russell. 2002. *Trust and Trustworthiness*. New York: Russell Sage Foundation.
- Hardin, Russell. 2004. "Internet Capital." *Analyse & Kritik* 26:122–138.
- Heath, Anthony. 1976. *Rational Choice and Social Exchange. A Critique of Exchange Theory*. Cambridge: Cambridge University Press.
- Heckathorn, Douglas D. 1989. "Collective Action and the Second-Order Free-Rider Problem." *Rationality and Society* 1:78–100.
- Jappelli, Tullio and Marco Pagano. 2002. "Information sharing, lending and defaults: Cross-country evidence." *Journal of Banking & Finance* 26:2017–2045.
- Jian, Lian, Jeffrey K. MacKie-Mason, and Paul Resnick. 2010. "I Scratched Yours: The Prevalence of Reciprocation in Feedback Provision on eBay." *The B.E. Journal of Economic Analysis & Policy* 10:Article 92.
- Jones, Candace, William S. Hesterly, and Stephen P. Borgatti. 1997. "A General Theory of Network Governance: Exchange Conditions and Social Mechanisms." *Academy of Management Review* 22:911–945.
- Kollock, Peter. 1994. "The Emergence of Exchange Structures: An Experimental Study of Uncertainty, Commitment, and Trust." *American Journal of Sociology* 100:313–345.
- Kollock, Peter. 1999. "The Production of Trust in Online Markets." *Advances in Group Processes* 16:99–123.
- Livingston, Jeffrey A. 2003. "What Attracts a Bidder to a Particular Internet Auction?" Pp. 165–187 in *Advances in Applied Microeconomics*, vol. 12, edited by M. R. Baye. Amsterdam: Elsevier.
- Lucking-Reiley, David, Doug Bryan, Naghi Prasad, and Daniel Reeves. 2007. "Pennies from eBay: The Determinants of Price in Online Auctions." *Journal of Industrial Economics* 55:223–233.

- Malinowski, Bronislaw. 1922. *Argonauts of the Western Pacific: An Account of Native Enterprise and Adventure in the Archipelagoes of Melanesian New Guinea*. London: Routledge & Kegan Paul.
- Malinowski, Bronislaw. 1926. *Crime and Custom in Savage Society*. London: Routledge & Kegan Paul.
- Mauss, Marcel. [1950] 1990. *The Gift: The Form and Reason for Exchange in Archaic Society*. London: Routledge.
- Milgrom, Paul R., Douglass C. North, and Barry R. Weingast. 1990. "The Role of Institutions in the Revival of Trade: The Law Merchant, Private Judges, and the Champagne Fairs." *Economics and Politics* 2:1–23.
- Nowak, Martin A. and Karl Sigmund. 2005. "Evolution of Indirect Reciprocity." *Nature* 437:1291–1298.
- Przepiorka, Wojtek. 2013. "Buyers pay for and sellers invest in a good reputation. More evidence from eBay." *Journal of Socio-Economics* 42:31–42.
- de Quervain, Dominique J.-F., Urs Fischbacher, Valerie Treyer, Melanie Schellhammer, Ulrich Schnyder, Alfred Buck, and Ernst Fehr. 2004. "The Neural Basis of Altruistic Punishment." *Science* 305:1254–1258.
- Resnick, Paul and Richard Zeckhauser. 2002. "Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System." Pp. 127–157 in *The Economics of the Internet and E-Commerce*. Vol. 11, *Advances in Applied Microeconomics*, edited by M. R. Baye. Amsterdam: Elsevier.
- Resnick, Paul, Richard Zeckhauser, Eric Friedman, and Ko Kuwabara. 2000. "Reputation Systems." *Communications of the ACM* 43:45–48.
- Resnick, Paul, Richard Zeckhauser, John Swanson, and Kate Lockwood. 2006. "The Value of Reputation on eBay: A Controlled Experiment." *Experimental Economics* 9:79–101.
- Schmieder, Johannes F. 2009. "gpreg: Stata Module to Estimate Regressions with Two Dimensional Fixed Effects." Boston College, Department of Economics, Statistical Software Components, S457048.
- Shapiro, Carl. 1983. "Premiums for High Quality Products as Return to Reputation." *Quarterly Journal of Economics* 98:659–680.
- Snijders, Chris and Jeroen Weesie. 2009. "Online Programming Markets." Pp. 166–185 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage Foundation.
- Trivers, Robert L. 1971. "The Evolution of Reciprocal Altruism." *Quarterly Review of Biology* 46:35–57.
- Uzzi, Brian. 1997. "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness." *Administrative Science Quarterly* 42:35–67.
- Yamagishi, Toshio, Karen S. Cook, and Motoki Watabe. 1998. "Uncertainty, Trust and Commitment Formation in the United States and Japan." *American Journal of Sociology* 104:165–194.
- Yamagishi, Toshio, Masafumi Matsuda, Noriaki Yoshikai, Hiroyuki Takahashi, and Yukihiro Usui. 2009. "Solving the Lemons Problem with Reputation." Pp. 73–108 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage Foundation.
- Zucker, Lynne G. 1986. "Production of Trust: Institutional Sources of Economic Structure, 1840–1920." *Research in Organizational Behavior* 8:53–111.

Figure 1. Survival functions and hazard rates of sellers' and buyers' rating decisions

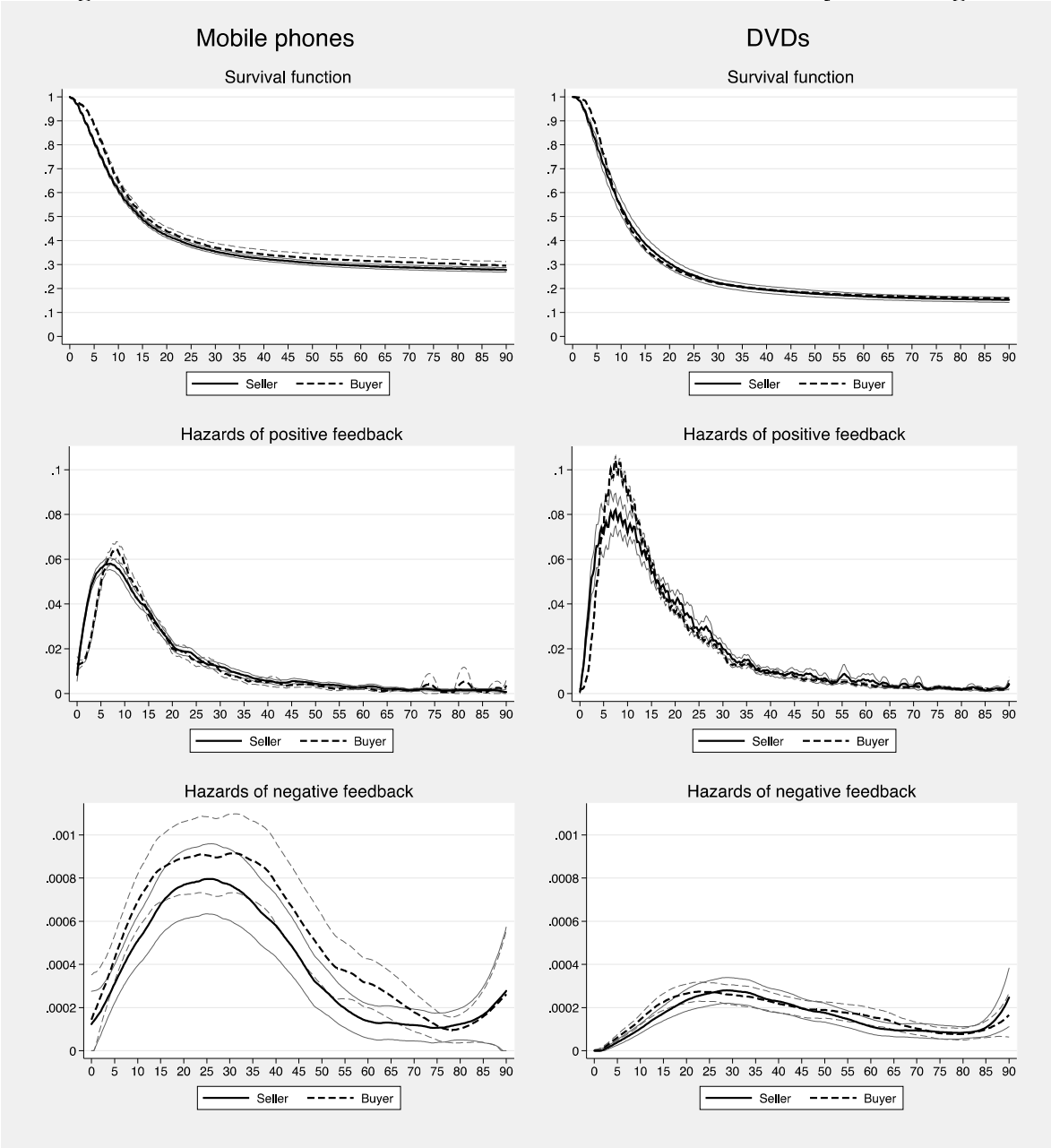


Table 1. Means and standard deviations of key variables

	New mobile phones	Used mobile phones	DVDs
Seller:			
Seller's positive ratings (log)	3.77 (1.59)	3.74 (1.61)	7.00 (2.65)
Seller's negative ratings (log)	.42 (.70)	.47 (.71)	2.00 (2.01)
Buyer (if sold):			
Buyer's positive ratings (log)	3.06 (1.70)	3.16 (1.83)	3.92 (1.69)
Buyer's negative ratings (log)	.27 (.53)	.32 (.62)	.28 (.54)
Feedback (if sold):			
Time until seller feedback (censored)	32.6 (35.3)	36.5 (36.7)	24.9 (30.1)
Positive seller feedback	.736	.685	.843
Neutral seller feedback	.002	.005	.001
Negative seller feedback	.012	.016	.003
Seller has previous rating from buyer	.032	.031	.049
Time until buyer feedback (censored)	33.9 (35.7)	39.0 (36.8)	25.2 (30.2)
Positive buyer feedback	.706	.656	.835
Neutral buyer feedback	.007	.011	.004
Negative buyer feedback	.019	.020	.004
Buyer has previous rating from seller	.030	.037	.049
Auction:			
Successful auction	.958	.956	.533
Selling price in Euros (if sold)	216.7 (37.4)	150.3 (47.1)	7.4 (10.2)
Number of observations	5499	9128	339517
Number of sellers	4341	7687	33166
Number of buyers	4881	7454	100046

Notes: Descriptive statistics of further control variables that are used in our model estimations are listed in Table A1 in the online supplement.

Table 2. Effect of reputation on sales and prices

	Product Sold (0/1)			Selling Price			Selling Price (with Fixed Effects)			
	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10
	New mobile phones	Used mobile phones	DVDs	New mobile phones	Used mobile phones	DVDs	New mobile phones	Used mobile phones	DVDs	DVDs
Seller's positive ratings (log)	.344* (.136)	-.019 (.064)	.117** (.040)	.005*** (.001)	.008* (.003)	.053*** (.015)	.027** (.010)	-.002 (.045)	.016*** (.002)	.005 (.007)
Seller's negative ratings (log)	-.670* (.301)	-.207 (.132)	-.145* (.058)	-.011*** (.002)	-.018** (.006)	-.101** (.035)	-.055 (.034)	-.025 (.033)	-.012** (.005)	-.036** (.014)
McFadden R-squared	.858	.678	.133							
R-Squared				.844	.513	.111				
Number of observations	5499	9128	339517	5269	8727	180881	1612	1944	113276	103030
Number of sellers	4341	7687	33166	4242	7474	30018	585	691		6901
Number of titles									18054	15964

Notes: M1-M3: Coefficient estimates of logistic regressions are shown; the binary dependent variable is equal to one for successful auctions. M4-M6: Coefficient estimates of OLS regressions are shown; the dependent variable is the logarithm of the selling price (in Euros). M7-M10: Coefficient estimates of fixed-effects regressions are shown (seller fixed effects in M7 and M8, title fixed effects in M9, title and seller fixed effects in M10); the dependent variable is the logarithm of the selling price (in Euros). Numbers in parentheses are robust standard errors (adjusted for seller-clusters; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ for two-sided tests). The models contained various control variables (starting price, length of product description, number of competing offers, and dummies for private profile, verified identity, Me-Page, PowerSeller, auction picture, thumbnail listing, bold listing, payment modes, auction duration, time and date of auction ending, and product subcategory); for detailed results see the online supplement.

Table 3. Feedback patterns (%)

	Mobile phones	DVDs
Auctions with mutual feedback:		
Buyer rated first	32.3	48.6
Seller rated first	28.1	30.2
Simultaneous	.1	.3
Auctions with buyer feedback only	10.1	5.1
Auctions with seller feedback only	11.8	5.7
Auctions without feedback	17.7	10.1
Total	100.0	100.0
Observations	13996	180881

Table 4. Hazards of positive and negative feedback in the DVD market

	Positive Feedback		Negative Feedback	
	M11	M12	M13	M14
	Buyer	Seller	Buyer	Seller
Positive first move by partner (time-dependent)	5.106*** (.112)	13.754*** (1.013)	.003 (.004)	-.006** (.002)
Neutral first move by partner (time-dependent)	.410 (.732)	2.176*** (.461)	-.011** (.004)	.440*** (.109)
Negative first move by partner (time-dependent)	-.544 (.327)	1.859*** (.461)	.535** (.166)	1.871*** (.311)
Positive ratings (log)	-.880*** (.183)	-.421 (.234)	-.025* (.012)	-.012** (.004)
Negative ratings (log)	-.616 (.508)	-.225 (.258)	-.102* (.051)	-.008 (.006)
Partner's positive ratings (log)	-.840*** (.137)	-.259 (.269)	.017*** (.005)	.021*** (.006)
Partner's negative ratings (log)	1.023*** (.259)	-.894 (.755)	-.043** (.014)	-.059 (.035)
Previous interaction rating:				
Received only	-.067 (.459)	.943 (.596)		
Provided only	-2.051*** (.428)	-1.448* (.676)		
Received and provided	-1.147*** (.219)	-.158 (.286)		
Received or provided			.002 (.005)	.006* (.003)
Number of observations	96055	96055	96055	96055
Number of events	80601	80343	310	223
Number of sellers	9309	9309	9309	9309
Number of buyers	26188	26188	26188	26188

Notes: The table shows coefficient estimates for the effects on the conditional probability of submitting a rating on a specific day given that no rating had been submitted yet (scaled by a factor of 100) for discrete-time linear probability models with seller and buyer fixed effects (standard errors in parentheses, adjusted for buyer-clusters in M11 and M13 and for seller-clusters in M12 and M14; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ for two-sided tests). Time indicators parameterizing the baseline hazard are not displayed.

Online Supplement for

Reputation Formation and the Evolution of Cooperation in Anonymous Online Markets

Andreas Diekmann^a, Ben Jann^b, Wojtek Przepiorka^{cd}, Stefan Wehrli^a

^a ETH Zurich, Sociology, CLU D, Clausiusstrasse 50, CH-8092 Zurich, Switzerland

^b University of Bern, Institute of Sociology, Lerchenweg 36, CH-3000 Bern 9, Switzerland

^c Nuffield College, New Road, Oxford OX1 1NF, UK

^d University of Oxford, Department of Sociology, Manor Road, Oxford OX1 3UQ, UK

Contents

Literature review	2
Details of data collection.....	4
Control variables included in the models in Table 2.....	4
Discussion of the effects of the control variables in the models in Table 2.....	5
Computation of hazard rates and confidence intervals in Figure 1	7
Model diagnostics	7
References	8
Table A1. Means and standard deviations of control variables	11
Table A2. Effect of reputation on sales.....	12
Table A3. Effect of reputation on price I	13
Table A4. Effect of reputation on price II	14

Literature review

There are several empirical studies investigating the effects of reputation in online markets (for earlier reviews, see Bajari and Hortacsu 2004; Resnick et al. 2006). Most of these studies analyze auction data collected from the online market platform eBay (for other platforms, see Ockenfels 2003; Snijders and Zijdemans 2004; Diekmann et al. 2009; Snijders and Weesie 2009) and use statistical models to estimate the effects of seller ratings on selling price and probability of sale. The studies employ data sets of different sizes (ranging from 100 to 340,000 cases) and covering various commodities (guitars, coins, stamps, mobile phones, DVDs, etc.). They differ in the definition of the dependent variable, the operationalization of sellers' reputation, and the control variables included in the models. The majority of the studies estimate the elasticity or semi-elasticity of the selling price with respect to the number of positive and negative ratings or with respect to the "score" (positive minus negative ratings) and account for censoring (i.e. that prices are unobserved for unsold items).

Some studies provide mixed results (Standifird 2001; Bajari and Hortacsu 2003; Lucking-Reiley 2007) or do not support the reputation hypothesis (Eaton 2005; Cabral and Hortacsu 2010). In the majority of studies, however, consistent evidence for a positive relation between sellers' reputation and selling price is found (McDonald and Slawson 2002; Melnik and Alm 2002; 2005; Dewan and Hsu 2004; Snijders and Zijdemans 2004; Berger and Schmitt 2005; Livingston 2005; Dewally and Ederington 2006; Houser and Wooders 2006; Przepiorka 2013). With respect to the probability of sale, the evidence is less conclusive. While the results of Dewan and Hsu (2004), Hsu, Hannan and Koçak (2009), and Przepiorka (2013) support the hypothesis of a positive (negative) relation between positive (negative) ratings and sales, Berger and Schmitt (2005), Eaton (2005), and Cabral and Hortacsu (2010) find no support or mixed evidence. Snijders and Zijdemans (2004) even find a negative relation between the score and probability of sale.¹

Apart from variables accounting for product characteristics (e.g. brands, book values, condition, etc.), the explanatory and control variables most frequently used are the starting price, the number of bids or bidders, the existence of a product picture, auction duration, whether or not the offer ends on a weekend, the availability of secure payment methods, and

¹ Snijders and Zijdemans (2004) and Eaton (2005) do not include the starting bid (i.e. minimum price) in their model. If the starting bid is negatively (positively) correlated with sellers' positive (negative) ratings (see, for example, Przepiorka 2013), the corresponding estimates of the reputation effects can be biased in the opposite direction.

whether or not the seller set a secret reserve price. While the starting price consistently and in line with expectations has a significantly negative effect on sales (Dewan and Hsu 2004; Berger and Schmitt 2005; Cabral and Hortacsu 2010; Przepiorka 2013), the merit of its inclusion in a model for the selling price is debatable. On the one hand, the selling price is left-censored at the starting price and, therefore, the effect of the starting price will be overestimated in a standard regression model. On the other hand, the starting price may have a genuine effect on the selling price. For example, a high starting price may be a signal of superior product quality or, conversely, a low starting price may attract more potential buyers (see Livingston 2003). Consequently, reputation effects can be biased in models that do not control for starting prices. The number of bids and the number of bidders are endogenous variables mediating the effects of reputation and should not be included in the model; the same is true for all buyer characteristics.

Providing a picture of the offered item is found to increase selling price by Snijders and Zijdemann (2004), Eaton (2005), and Dewally and Ederington (2006), as well as the probability of sale by Snijders and Zijdemann (2004). However, Eaton (2005) and Cabral and Hortacsu (2010) report a decreasing effect on the probability of sale and Melnik and Alm (2002; 2005) did not find any effects whatsoever in regard to the product picture. Auction duration is found to have no effect (Melnik and Alm 2002; 2005; Houser and Wooders 2006) or a positive effect on sales (Dewan and Hsu 2004; Cabral and Hortacsu 2010; Przepiorka 2013) and prices (Dewally and Ederington 2006; Lucking-Reiley et al. 2007; Cabral and Hortacsu 2010; Przepiorka 2013). Results with respect to an auction ending on a weekend are mixed (Melnik and Alm 2002; 2005; Dewan and Hsu 2004; Lucking-Reiley et al. 2007; Przepiorka 2013). Moreover, with the exception of two studies (Cabral and Hortacsu 2010; Przepiorka 2013), there is no evidence that the availability of a secure payment method positively affects sales or prices (Melnik and Alm 2002; 2005; Berger and Schmitt 2005; Houser and Wooders 2006). Finally, secret reserve prices are commonly used in online auction markets in the US or the UK, but on eBay Germany, in most product categories, sellers do not have a possibility to set secret reserve prices. If sellers set a secret reserve price, a transaction takes place only if the highest bid reaches that threshold. While two studies find a positive and significant effect of a secret reserve price on selling price (Dewan and Hsu 2004; Lucking-Reiley et al. 2007), results are mixed in one study (Standifird 2001) and two further studies find a significantly negative effect (Bajari and Hortacsu 2003; Dewally and Ederington 2006).

Details of data collection

We collected data on a total of 1,084,882 auctions of various products on the German eBay market platform between November 7, 2004 and January 7, 2005. Information on auction characteristics was gathered using a computer program written to observe auctions for scientific analysis. Based on a set of 144 eBay categories (e.g. “Sony-Ericsson T610”), the first program maintained a directory of identification numbers and prospective end dates of new auctions that was updated daily. The auctions in this list were then processed sequentially by a second program, which scanned an auction’s pages once the end date was reached. The following pages were accessed and stored in a local database: (1) the auction’s main page (the “item page”); (2) the list of bids; (3) the profile page of the seller; (4) the profile page of the buyer; (5) the page listing other items being offered by the seller; and (6) the page listing offers by the buyer. The information on these pages was extracted using regular expressions and stored in database fields. For the item page, the HTML source code was also stored. After 90 days, the maximum duration for submitting feedback, a third program recorded the entire feedback history of the seller and the buyer involved in the transaction.

The usual procedure on eBay is to handle transactions in the manner of an English auction, but the system also allows sellers to set a so-called “Buy It Now” price at which the product can be purchased immediately. We recorded data on offers where the English auction mechanism was used, possibly in combination with a Buy It Now price, but disregarded offers in which the seller disabled the English auction mechanism and employed a Buy It Now price only.

Due to an error, only partial data was recorded for auctions with more than 18 bids. We tried to recover the missing data after detecting the problem, but this was only possible for auctions that ended on December 1, 2004 or later. Auctions that had ended earlier were taken offline and we could not retrieve them using the item number. To have complete data and a well-defined sample, we therefore restricted our analysis to auctions that ended between December 1, 2004 and January 7, 2005.

Control variables included in the models in Table 2

Apart from positive and negative ratings, we controlled for other seller characteristics, such as whether they had a private profile, a verified identity, a Me-Page or PowerSeller status.²

² A private profile means that, although rating statistics are visible, text feedback is hidden from other users (and “private” is displayed next to the user’s name instead of the feedback score); a verified identity can be obtained

These attributes are more common for sellers in the DVD market than for sellers in the mobile phone market, which indicates that professionalism is more pronounced in the latter (see Table A1).

In addition to the product subcategories (*model* for mobile phones, *genre* for DVDs), control variables included the starting price (minimum bid; models M1–M3 only), the duration of the auction, the payment modes offered, the (logarithm of the) number of characters used in the product description (as a proxy for the amount of information provided about the offered item), whether a picture was provided, and whether the auction was highlighted in eBay’s list of offers using boldface or a picture thumbnail (eBay charges a small fee for each of these options). We further included a variable called “competition” that captures the (logarithm of the) number of similar offers that were active at the time the auction ended (same model and same new/used status in the case of mobile phones and same genre in the case of DVDs).

Finally we included variables for the date and time the auction ended. In our models we used dummy variables for the week (the observation window covers six weeks), the weekday (plus a dummy for December 24/25 or December 31/January 1), and the daytime (one eight-hour interval starting at midnight followed by four-hour intervals in the case of mobile phones; two four-hour intervals starting at midnight and at 4 a.m. followed by hourly intervals in the case of DVDs).

Unfortunately, we were not able to collect reliable information on shipping costs, which are usually charged by the seller in addition to the selling price. We therefore did not include such a variable in our models. Furthermore, we did not include the number of bids or buyer characteristics as explanatory variables because these variables are endogenous.

Discussion of the effects of the control variables in the models in Table 2

Results for the full models for the probability of sale (M1–M3) are displayed in Table A2. Regarding seller characteristics other than positive and negative ratings, no clear pattern can be found. The large negative effect of a seller’s verified identity in model M1 may be an artifact driven by the small number of sellers with this attribute. Not surprisingly, the starting price is a strong determinant of sales in all models. The higher the starting price, the less likely it is that the item will be sold.

for a small fee by proving one’s identity at a post office; a Me-Page is a user’s homepage on eBay that can be used to provide additional information on a business or person; and PowerSeller status is granted to professional sellers with a certain trade volume and a good feedback record.

As expected, the level of competition (the number of similar offers currently on the market) exerts a negative effect on the chances of selling a product (negative effect in all models, but not significant for new mobiles). The options to increase the visibility of an auction in the list of offers (thumbnail, boldface) seem to serve their purpose, but their effects are only significant in the model for DVDs. The description length has the expected positive effect for (new) mobile phones, but negatively affects DVD sales. An *ad hoc* explanation for this counter-intuitive finding could be that an extensive description is valuable for complex goods such as mobile phones, but is irrelevant or might even raise suspicion in the case of an item such as a DVD. Furthermore, not providing a picture has a negative effect for (used) mobile phones, but does not affect the sales of DVDs. The models also contain a number of dummy variables to control for the payment modes offered (bank transfer, PayPal or credit card, cash on pick up, cash on delivery), auction duration (one, three, five, seven or 10 days), the time-point of the end of the auction (using a total of 16 and 29 parameters for mobile phones and DVDs, respectively; see the description of the control variables above), and the product subcategory (seven mobile phone models and seven DVD genres). For these groups of variables, joint tests are reported in the lower panel of the table without listing the single coefficients. The product subcategory is highly significant in all models, the auction duration is significant for DVDs but not for mobile phones, and the time-point of the end of the auction is significant for DVDs and used mobile phones but not for new mobile phones.

Results for the full models for the selling price (M4–M10) are displayed in Table A3 and in Table A4. Positive and negative ratings are not the only source of information a buyer can use to assess a seller and his or her products. Auction characteristics such as a picture and the description length also have a positive effect on price for both new and used mobile phones. We also observe a positive effect of the description length for DVDs (counteracting the negative effect on the probability of sale), but providing a picture, again, seems to be less important for DVDs. Seller characteristics such as a verified identity, a Me-Page or PowerSeller status do not exert a positive effect on price. A likely reason for the lack of an effect of PowerSeller status is that PowerSeller status can only be acquired by sellers who already have a high number and proportion of positive ratings. Because of the decreasing marginal effect of positive information, PowerSeller status does not contribute much to increase a seller's trustworthiness and therefor is discarded by potential buyers. Moreover, sellers with a private profile do not seem to be disadvantaged. This might appear surprising,

but recall that a private profile does not hide the seller's feedback statistics but only the corresponding text comments.

In accordance with theoretical considerations, more intense competition lowers the price for mobile phones. For DVDs, no competition effect can be found, but this is not very surprising given the broad nature of the competition measure in this case (recall, however, that competition had an effect on sales of DVDs; see Table A2). Similarly, increasing the visibility of an offer using a thumbnail or bold listing seems to attract bidders and, as expected, increase the price. Payment modes offered take up some of the unexplained variance except for used mobile phones. The auction duration does not have an impact on price, but there are significant timing effects. In the case of mobile phones, the model makes a large difference and there are also some price differences among DVD genres.

We do not include the starting price in Models M4–M10 because the selling price is censored at the starting price (causing the coefficient for the starting price to be positively biased). Consistent estimates for the effect of the starting price can be obtained via censored regression. In our case, the starting price exerts a weak positive effect for used mobile phones and DVDs, but not for new mobile phones (not shown).

Computation of hazard rates and confidence intervals in Figure 1

The displayed hazard rates are kernel-density estimates based on nonparametric hazard components, applying linear boundary correction at days 0 and 90 (Jones 1993) and using a direct plug-in automatic bandwidth (Wand and Jones 1995), as implemented by Jann (2005). The dotted lines are normal approximation bootstrap confidence intervals taking into account clustering of sellers and buyers, respectively.

Model diagnostics

All reported results are fairly robust against changes in model specifications. For example, the exact choice of control variables does not affect the effects of key variables much. Because a selling price cannot be smaller than the starting price we also replicated the results for the selling price (M4–M6) using models for censored data (censored normal regression, censored normal regression with heteroscedastic errors, Heckman selection regression, censored quantile regression). In all variants the effects of the reputation variables were highly significant. Furthermore, the functional forms of effects of continuous variables were evaluated using graphical methods. No major deviations from the chosen forms in the

reported models were detected. Multicollinearity does not seem to be a problem either. The only variable that reaches a substantial value for the variance inflation factor is our measure for competition, which strongly depends on the auction date and time. However, excluding competition from the models does not change our findings. Finally, the results do not seem to be driven by outliers. We replicated the models excluding the most influential percent of observations (as measured by Cook's D) without changing our substantial conclusions.

Three additional issues are worth mentioning with respect to discrete-time linear probability models (LPM) we used to analyze the feedback behavior (M11–M14). First, an alternative method of analyzing the feedback process would be to apply a proportional-hazards model for continuous time-to-event data (Cox regression; see e.g. Kalbfleisch and Prentice 1980). We did not employ this model because there is no satisfactory way to incorporate buyer and seller fixed effects simultaneously (for the same reason, we use an LPM instead of the more common discrete-time logit or complementary log-log model). An analysis of our data using fixed effects (i.e. stratified) proportional-hazards models (Allison 2009), where each model is estimated twice, once stratified by sellers and once by buyers, leads to substantially the same conclusions. Second, in our discrete-time LPMs we did not include auction characteristics (e.g. the selling price) as control variables. Including such variables does not change our results. Third, in our feedback models, effects are assumed to be constant over process time. This assumption seems to be violated for the reciprocity effect (i.e. the effect of a first move by the partner). Detailed analyses allowing the reciprocity effect to vary by process time indicate that the effect is strongest in the beginning, but stays positive over the whole 90-day period.

References

- Allison, Paul D. 2009. *Fixed Effects Regression Models*. Thousand Oaks, CA: Sage.
- Bajari, Patrick and Ali Hortacsu. 2003. "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions." *RAND Journal of Economics* 34:329–355.
- Bajari, Patrick and Ali Hortacsu. 2004. "Economic Insights from Internet Auctions." *Journal of Economic Literature* 42:457–486.
- Berger, Roger and Katharina Schmitt. 2005. "Vertrauen bei Internetauktionen und die Rolle von Reputation, Informationen, Treuhandangebot und Preisniveau." *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 57:86–111.
- Cabral, Luis M.B. and Ali Hortacsu. 2010. "The Dynamics of Seller Reputation: Evidence from eBay." *Journal of Industrial Economics* 58:54–78.
- Dewally, Michaël and Louis Ederington. 2006. "Reputation, Certification, Warranties, and Information as Remedies for Seller-Buyer Information Asymmetries: Lessons from the Online Comic Book Market." *Journal of Business* 79:693–729.
- Dewan, Sanjeev and Vernon Hsu. 2004. "Adverse Selection in Electronic Markets: Evidence from Online Stamp Auctions." *Journal of Industrial Economics* 52:497–516.
- Diekmann, Andreas, Ben Jann, and David Wyder. 2009. "Trust and Reputation in Internet Auctions." Pp. 139–165 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage.
- Eaton, David H. 2005. "Valuing Information: Evidence from Guitar Auctions on eBay." *Journal of Applied Economics and Policy* 24:1–19.
- Houser, Daniel and John Wooders. 2006. "Reputation in Auctions: Theory, and Evidence from eBay." *Journal of Economics & Management Strategy* 15:353–369.
- Hsu, Greta, Michael T. Hannan, and Özgecan Koçak. 2009. "Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests." *American Sociological Review* 74:150–169.
- Jann, Ben. 2005. "kdens: Stata Module for Univariate Kernel Density Estimation." Boston College, Department of Economics, Statistical Software Components, S456410.
- Jones, M. C. 1993. "Simple Boundary Correction for Kernel Density Estimation." *Statistics and Computing* 3:135–146.
- Kalbfleisch, John. D. and Ross L. Prentice. 1980. *The Statistical Analysis of Failure Time Data*. New York: John Wiley & Sons.
- Livingston, Jeffrey A. 2003. "What Attracts a Bidder to a Particular Internet Auction?" Pp. 165–187 in *Organizing the New Industrial Economy*. Vol. 12, *Advances in Applied Microeconomics*, edited by M. R. Baye. Amsterdam: Elsevier.
- Livingston, Jeffrey A. 2005. "How Valuable is a Good Reputation? A sample selection model of Internet auctions." *Review of Economics and Statistics* 87:453–465.
- Lucking-Reiley, David, Doug Bryan, Naghi Prasad, and Daniel Reeves. 2007. "Pennies from eBay: The Determinants of Price in Online Auctions." *Journal of Industrial Economics* 55:223–233.
- McDonald, Cynthia G. and Carlos Jr. V. Slawson. 2002. "Reputation in an Internet Auction Market." *Economic Inquiry* 40:633–650.
- Melnik, Mikhail I. and James Alm. 2002. "Does a Seller's Ecommerce Reputation Matter? Evidence from eBay Auctions." *Journal of Industrial Economics* 50:337–349.
- Melnik, Mikhail I. and James Alm. 2005. "Seller Reputation, Information Signals, and Prices for Heterogeneous Coins on eBay." *Southern Economic Journal* 72:305–328.
- Ockenfels, Axel. 2003. "Reputationsmechanismen auf Internet-Marktplattformen." *Zeitschrift für Betriebswirtschaft* 73:295–315.

- Przepiorka, Wojtek. 2013. "Buyers pay for and sellers invest in a good reputation. More evidence from eBay." *Journal of Socio-Economics* 42:31–42.
- Resnick, Paul, Richard Zeckhauser, John Swanson, and Kate Lockwood. 2006. "The Value of Reputation on eBay: A controlled experiment." *Experimental Economics* 9:79–101.
- Snijders, Chris and Jeroen Weesie. 2009. "Online Programming Markets." Pp. 166–185 in *eTrust: Forming Relationships in the Online World*, edited by K. S. Cook, C. Snijders, V. Buskens, and C. Cheshire. New York: Russell Sage Foundation.
- Snijders, Chris and Richard Zijdemann. 2004. "Reputation and Internet Auctions: eBay and Beyond." *Analyse & Kritik* 26:158–184.
- Standifird, Stephen S. 2001. "Reputation and e-commerce: eBay auctions and the asymmetrical impact of positive and negative ratings." *Journal of Management* 27:279–295.
- Wand, M. P. and M. C. Jones. 1995. *Kernel Smoothing*. London: Chapman and Hall.

Table A1. Means and standard deviations of control variables

	New mobile phones	Used mobile phones	DVDs
Seller:			
Seller has private profile	.025	.028	.017
Seller has verified identity	.032	.038	.188
Seller has Me-Page	.086	.091	.486
Seller is PowerSeller	.016	.020	.282
Buyer (if sold):			
Buyer has private profile	.021	.049	.008
Buyer has verified identity	.015	.014	.020
Buyer has Me-Page	.045	.050	.047
Buyer is PowerSeller	.007	.005	.002
Auction:			
Starting price in Euros	28.8 (69.7)	25.9 (57.0)	3.7 (6.6)
No picture	.050	.058	.071
Description length (log)	7.30 (1.15)	6.95 (1.12)	7.04 (1.36)
Competition (log)	3.65 (1.08)	3.87 (.83)	6.92 (.87)
Listed with thumbnail	.419	.352	.033
Listed in bold	.146	.118	.003
Payment modes offered:			
No bank transfer	.035	.034	.082
PayPal or credit card	.107	.082	.169
Cash on pick up	.479	.399	.169
Cash on delivery	.117	.087	.023
Auction duration:			
1 day	.142	.104	.291
3 days	.319	.278	.195
5 days	.191	.183	.112
7 days	.267	.344	.236
10 days	.081	.091	.166
Model (mobile) or Genre (DVDs):			
Sony T610 / Action	.041	.111	.403
Sony T630 / Drama	.061	.085	.061
Nokia 6230 / Classics	.531	.152	.051
Nokia 6310i / Comedy	.061	.296	.185
Motorola V600 / Crime	.089	.068	.093
Samsung E800 / Cartoon	.156	.062	.021
Samsung E700 / Children	.062	.226	.186
Number of observations	5499	9128	339517
Number of sellers	4341	7687	33166
Number of buyers	4881	7454	100046

Table A2. Effect of reputation on sales

	M1		M2		M3	
	New mobile phones		Used mobile phones		DVDs	
Seller's positive ratings (log)	.344 [*]	(.136)	-.019	(.064)	.117 ^{**}	(.040)
Seller's negative ratings (log)	-.670 [*]	(.301)	-.207	(.132)	-.145 [*]	(.058)
Seller has private profile	-.476	(.865)	-.532	(.593)	-.583 [*]	(.292)
Seller has verified identity	-3.218 ^{***}	(.712)	.235	(.366)	.085	(.206)
Seller has Me-Page	-.185	(.602)	-.100	(.327)	-.724 ^{***}	(.153)
Seller is PowerSeller	-.024	(.775)	-.126	(.419)	-.203	(.165)
Starting price	-.110 ^{***}	(.010)	-.056 ^{***}	(.005)	-.109 ^{***}	(.028)
No picture	-.883	(.803)	-.594 [*]	(.276)	.009	(.155)
Description length (log)	.423 ^{**}	(.159)	.047	(.078)	-.115 ^{***}	(.028)
Competition (log)	-.880	(.677)	-1.076 ^{***}	(.279)	-.656 ^{***}	(.111)
Listed with thumbnail	.026	(.385)	.280	(.175)	.692 ^{***}	(.150)
Listed in bold	1.074	(.679)	.432	(.289)	.731 [*]	(.307)
Constant	25.165 ^{***}	(3.686)	17.835 ^{***}	(2.018)	7.429 ^{***}	(.983)
Further covariates (Chi ² -tests):						
Payment modes (4 df)	9.7 [*]		2.5		16.1 ^{**}	
Auction duration (4 df)	4.1		7.5		36.8 ^{***}	
End of auction (16/29 df)	22.4		36.5 ^{**}		298.4 ^{***}	
Product subcategory (6 df)	82.8 ^{***}		113.3 ^{***}		65.5 ^{***}	
McFadden R-squared	.858		.678		.133	
Number of observations	5499		9128		339517	
Number of sellers	4341		7687		33166	

Note: The table shows coefficient estimates of logistic regressions and robust standard errors in parentheses (adjusted for seller-clusters; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ for two-sided tests). The binary dependent variable in all models is equal to one for successful auctions.

Table A3. Effect of reputation on price I

	M4		M5		M6	
	New mobile phones		Used mobile phones		DVDs	
Seller's positive ratings (log)	.005***	(.001)	.008*	(.003)	.053***	(.015)
Seller's negative ratings (log)	-.011***	(.002)	-.018**	(.006)	-.101**	(.035)
Seller has private profile	-.019	(.011)	.018	(.020)	-.027	(.103)
Seller has verified identity	.001	(.007)	-.002	(.019)	.107	(.154)
Seller has Me-Page	.007	(.005)	-.019	(.016)	-.267***	(.076)
Seller is PowerSeller	.006	(.012)	.000	(.034)	.020	(.118)
No picture	-.024**	(.009)	-.044*	(.022)	.067	(.089)
Description length (log)	.004***	(.001)	.026***	(.004)	.026*	(.011)
Competition (log)	-.018**	(.006)	-.037*	(.016)	-.022	(.027)
Listed with thumbnail	.010***	(.003)	.036***	(.007)	.321***	(.052)
Listed in bold	.006	(.004)	.032**	(.010)	.521***	(.081)
Constant	5.473***	(.031)	5.261***	(.085)	1.519***	(.244)
Further covariates (F-tests):						
Payment modes (4 df)	5.4***		0.7		3.8**	
Auction duration (4 df)	.4		1.3		.5	
End of auction (16/29 df)	26.6***		7.6***		4.6***	
Product subcategory (6 df)	1545.4***		979.8***		18.6***	
R-squared	.844		.513		.111	
Number of observations	5269		8727		180881	
Number of sellers	4242		7474		30018	

Note: The table shows coefficient estimates of OLS regressions and robust standard errors in parentheses (adjusted for seller-clusters; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ for two-sided tests). The dependent variable in all models is the logarithm of the selling price (in Euros).

Table A4. Effect of reputation on price II

	M7	M8	M9	M10
	New mobile phones	Used mobile phones	DVDs	DVDs
	Seller FEs	Seller FEs	Title FEs	Title and seller FEs
Seller's positive ratings (log)	.027** (.010)	-.002 (.045)	.016*** (.002)	.005 (.007)
Seller's negative ratings (log)	-.055 (.034)	-.025 (.033)	-.012** (.005)	-.036** (.014)
Seller has private profile			-.002 (.026)	.008 (.082)
Seller has verified identity			.011 (.013)	-.049* (.021)
Seller has Me-Page			.006 (.009)	-.083 (.047)
Seller is PowerSeller			-.024 (.013)	.000 (.015)
No picture	-.029* (.015)	-.063 (.073)	-.064*** (.010)	-.007 (.016)
Description length (log)	.003 (.006)	.041 (.031)	.009*** (.002)	-.004 (.005)
Competition (log)	.008 (.010)	-.037 (.028)	-.010* (.004)	-.002 (.004)
Listed with thumbnail	-.008 (.014)	.051 (.034)	.015 (.009)	.022* (.010)
Listed in bold	.018 (.015)	-.014 (.037)	.071 (.045)	-.009 (.072)
Constant	5.297*** (.084)	5.139*** (.308)	1.492*** (.037)	
Further covariates (F-tests):				
Payment modes (4 df)	1.5	.6	.9	1.4
Auction duration (4 df)	.4	1.6	5.7***	4.7***
End of auction (16/29 df)	7.3***	4.2***	8.2***	8.0***
Product subcategory (6 df)	55.0***	26.2***		
Number of observations	1612	1944	113276	103030
Number of titles			18054	15964
Number of sellers	585	691		6901

Note: The table shows coefficient estimates of fixed-effects (FE) regressions and robust standard errors in parentheses (adjusted for seller-clusters; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ for two-sided tests). The dependent variable in all models is the logarithm of the selling price (in Euros).